

# Sequence Analysis of Hepatitis C Virus Variants Producing Discrepant Results With Two Different Genotyping Assays

L.E. Prescott,<sup>1</sup> A. Berger,<sup>2</sup> J.-M. Pawlotsky,<sup>3</sup> P. Conjeevaram,<sup>2</sup> I. Pike,<sup>4</sup> and P. Simmonds<sup>1\*</sup>

<sup>1</sup>Department of Medical Microbiology, University of Edinburgh, Edinburgh, Scotland, UK

<sup>2</sup>Institut für Medizinische Virologie der J.W. Goethe-Universität Frankfurt, Frankfurt, Germany

<sup>3</sup>Department of Bacteriology and Virology, Hôpital Henri Mondor, Université de Paris, Créteil, France

<sup>4</sup>Murex Biotech Ltd, Dartford, Kent, UK

Methods for identifying the genotype of hepatitis C virus (HCV) in clinical specimens are frequently based upon the direct characterisation of viral RNA sequences by polymerase chain reaction (PCR) amplification, or by serologically based methods, in which the infecting genotype is inferred from the pattern of antibody reactivity to type-specific peptides or recombinant proteins used as antigens in an Enzyme Linked Immunosorbent Assay (ELISA). Although genotyping by direct, PCR-based methods show generally highly concordant results with the genotype inferred from serological typing assays (>95% agreement), there exist a small number of samples that produce discrepant results. To investigate the underlying reasons for the discrepancies, we obtained eleven samples from haemophiliacs and four samples from patients with chronic hepatitis C that produced discordant results between a PCR based assay (InnoLipa I and II) and a serotyping assay (Murex HC02). Nucleotide sequences in the 5' noncoding region (5'NCR), core, and NS4 region were used to identify the genotype of the circulating virus and to identify amino acid changes in NS4 that might alter antigenicity. In 14 samples, sequence analysis of all three regions was concordant with the results of the InnoLipa assay. There were few if any amino acid substitutions in NS4 that might have accounted for the discrepant serotyping results, which were found predominantly in samples from individuals with a history of multiple exposure to HCV. It remains unclear whether the detection of antibody in such discrepant samples corresponds to previous expression of a different genotype than detected by PCR, or whether the virus population in plasma is more restricted in genotype diversity than the population in the liver or at other sites of viral replication. *J. Med. Virol.* 53:237-244, 1997.

© 1997 Wiley-Liss, Inc.

**KEY WORDS:** serotyping; genotyping; genotype; NS4; 5' noncoding region

## INTRODUCTION

Hepatitis C virus (HCV) is a positive strand RNA virus which was first cloned in 1989 and is the major aetiological agent of non-A non-B hepatitis worldwide [Choo et al., 1989]. The virus genome is approximately 9400 nucleotides in length, and displays up to 30% sequence divergence between each of the six major types, each of which can be further divided into numerous subtypes [Simmonds et al., 1993a; Bukh et al., 1993a; Simmonds et al., 1994a; Mellor et al., 1995a].

Recent investigations have suggested that infections with HCV type 1 are associated with more rapid disease progression and are more likely to cause hepatocellular carcinoma than infection with other genotypes such as types 2 or 3 [Takada et al., 1996; Kobayashi et al., 1996]. Another reason for investigating the infecting HCV genotype in clinical samples is in the evaluation of patients for antiviral treatment, since there is an increased frequency of sustained response to interferon in individuals infected with types 2 and 3, compared to HCV type 1 [reviewed in Davis, 1994; Dushenko et al., 1994].

Because of these clinical associations, a number of different methods have been developed to identify the virus genotypes in infected individuals. The majority of techniques are based on the amplification of subgenomic fragments of HCV by the polymerase chain reaction (PCR). Genotypes can be distinguished using amplified DNA sequences in three different ways: (1) by differential amplification using type specific primers [Okamoto et al., 1992; Okamoto et al., 1993; Widell et

\*Correspondence to: P. Simmonds, Department of Medical Microbiology, University of Edinburgh, Teviot Place, EH8 9AG, United Kingdom.

Accepted 11 June 1997

al., 1994; Ohno et al., 1997], (2) by the digestion of amplified PCR products from the 5' noncoding region (5'NCR) with combinations of restriction endonucleases, to reveal patterns of fragments specific to individual genotypes [McOmish et al., 1994; Davidson et al., 1995; Mellor et al., 1996], or (3) by the reverse hybridisation of labelled PCR products to immobilised type specific probes [Stuyver et al., 1993; Stuyver et al., 1995; Stuyver et al., 1996]. A different, indirect method for typing virus in infected individuals is based upon the observation that sequence diversity between genotypes results in the production of type-specific antibodies, which can be detected by ELISA [Machida et al., 1992; Tanaka et al., 1994; Simmonds et al., 1993; Zhang et al., 1995; Dixit et al., 1995; Bhattacharjee et al., 1995]. The usefulness of serotyping assays lies in their rapidity and ease of use, as long as this is not at the expense of sensitivity or accuracy.

In this study, we have compared the performance of one direct (Inno-LiPA I and II) and one indirect (Murex 1-6 serotyping) typing assay on two distinct populations: a group of haemophiliacs from Germany, and patients with chronic hepatitis C from France. Discrepant samples were further investigated by sequence analysis of the 5'NCR, core and NS4 regions.

## MATERIALS AND METHODS

### Samples

Serum samples were obtained from haemophiliacs in Frankfurt, Germany, and from patients with chronic HCV infection in France. Additional samples were obtained from some individuals with discrepant results in order to permit sequence analysis.

### Genotyping

Isolates were genotyped using Inno-LiPA HCV version I or version II (Innogenetics, Antwerp, Belgium) according to the manufacturer's instructions. This assay used labelled PCR products from the 5'NCR as probes for reverse hybridisation to oligonucleotides immobilised on strips [Stuyver et al., 1993; Stuyver et al., 1996].

### Serotyping

Type-specific antibody was determined by the Murex 1-6 serotyping assay HC02 (Murex Biotech, Dartford, UK), or where specified, the Murex 1-3 serotyping assay HC01, according to the manufacturer's instructions. Type-specific peptides corresponding to two epitopes in NS4 were used as antigen in a competitive ELISA [Simmonds et al., 1993; Bhattacharjee et al., 1995].

### RNA Extraction and Nested Reverse-Transcription PCR (RT-PCR)

Virus RNA was directly extracted from 100  $\mu$ l serum using proteinase K/sarcosyl and phenol/chloroform extraction. RNA was precipitated at  $-20^{\circ}\text{C}$  overnight in ethanol and dissolved in 25  $\mu$ l DEPC treated water prior to RT-PCR [Davidson et al., 1995].

**(1) 5' NCR.** Reverse transcription of RNA used the outer antisense primer 209, followed by primary PCR with primers 209 and 939 (outer sense) and secondary PCR with inner primers 940 (sense) and 211 (antisense) [Chan et al., 1992].

**(2) Core.** Core specific antisense primer 410 was used for reverse transcription of RNA, followed by PCR amplification with primers 410 and 954 (outer sense), and inner primers 951 (sense) and 953 (antisense) [Mellor et al., 1995].

**(3) NS4.** NS4 amplification was achieved using combinations of universal primers, and type-specific primers designed for HCV types 1, 2 and 3, as follows; outer antisense primers included 007 [Simmonds et al., 1993b], LP2 (type 1 specific) 5'GTACTGTATCCCGCTGATGAARTTCCACA 3', and LP4 (type 3 specific) 5'GTATTGGATCCCACTCACAAAATTCCACA 3'. Outer sense primers were 5668 5'ATGCATGTCRGCTGAYCTGGA 3' and 53510 [Bhattacharjee et al., 1995]. For the second round of PCR, primers were 220 [Simmonds et al., 1993b], 865 [Prescott et al., 1996], LP3 (antisense, type 1) 5'TTCCACATGTGYTTCGCCAGCA 3', LP6 (sense, type 1) 5'CTGGAGGTCGTCACCTAGCACCTGG 3', LP15 (antisense, type 2) 5'CAAAGCTTANAGCATCTCSGCNATCCGCTGCCC 3', LP12 (sense, type 2) 5'GCTCTCGAGCCSGACAAGGARRTCYTNTAT 3', LP5 (antisense, type 3) 5'TTCCACATATGCTTGTGCCAGA 3' and LP7 (sense, type 3) 5'CTGGAAGTAACCACCAGCACCTGG 3'. PCR products were electrophoresed through a 2% agarose gel containing 0.5  $\mu$ l/ml ethidium bromide, and viewed on a u.v. transilluminator.

### Sequencing

PCR products were either sequenced directly (5'NCR and core regions) or from clones. Direct sequencing involved the re-amplification of the 2 $^{\circ}$  PCR product with one of the primers biotinylated, followed by immobilisation on streptavidin-coated magnetic beads (Dynabeads M280, Dynal) as previously described [Kleter et al., 1994]. Sequencing reactions were performed using T7 DNA polymerase (Sequenase version 2.0, United States Biologicals, USB) according to the manufacturer's instructions. Sequences from the 5'NCR, core, and NS4 regions have been submitted to GenBank and bear the accession numbers AF007498-AF007522.

### Cloning of NS4

TA cloning of amplified NS4 fragments into the pTAG vector was carried out using the LigA-Tor kit (R&D systems, Europe, Ltd.) according to the manufacturer's instructions. Cloned NS4 fragments were sequenced after the denaturation of miniprep DNA with 2M NaOH/2mM EDTA, followed by precipitation in ethanol at  $-70^{\circ}\text{C}$  for 15 minutes. To investigate the possibility of a mixed infection, up to five different clones were sequenced for samples GD1, GD3, GD4, GD5, and GD7. Between five and ten individual clones were sequenced for each PCR product from samples giving positive PCR results with more than one set of

		-160	-140	-120	-100
1a HCV-PT	ATTGCCAGGACGACCGGGTCCTTTC--TTGGAT-CAACCCGCTCAATGCCTGGAGATTGGGCGTGCCCCGCAAGACTGCTA				
1a HC-J1	.....-A.....C.....				
1b HCV-J	.....-.....G.....				
1b BK	.....-.....G.....				
GD 6	.....-.....G.....				
FD 1	.....-.....				
1c HC-J9	.....-T.....				
2a HC-J6	.....G...A...T.....-A...A...T...C...TC				
2a K2a	.....G...A...T.....-A...A...T...C...TC				
2b HC-J8	...A..G..A..A...T.....-A...A...T...T.C...TC.....AC.....				
2b K2b	...A..G..A..A...T.....-A...A...T...T.C...TC.....				
GD 10	...A..G..A..A...T.....-A...A...T...T.C...TC.....				
GD 11	...A..G..A..A...T.....-A...A...T...T.C...TC.....				
2c BEBEI	.....G...A...T.....-A...A...T...C...TC				
2c T983	.....G...A...T.....-A...A...T...C...CC				
3a NZL1	..C..TG..GT.....-G-.....A..CA..A.....G...TCA...				
3a HPCBK6	..C..TG..GT.....-A-.....A..CA..A.....G...TCA...				
GD 1	..C..TG..GT.....-G-T.....A..CA..A.....G...TCA...				
GD 2	..C..TG..GT.....-G-T.....A..CA..A.....G...TCA...				
GD 3	..C..TG..GT.....-G-.....A..CA..A.....G...TCA...				
GD 4	..C..TG..GT.....-G-T.....A..CA..A.....G...TCA...				
GD 5	..C..TG..GT.....-G-.....A..CA..A.....G...TCA...				
GD 7	..C..TG..GT.....-A-.....A..CA..A.....G...TCA...				
GD 8	..C..TG..GT.....-A-.....A..CA..A.....G...TCA...				
GD 9	..C..TG..GT.....-G-.....A..CA..A.....G...TCA...				
FD 3	..C..TG..GT.....-G-.....A..CA..A.....G...TCA...				
FD 4	..C..TG..GT.....-A-.....A..CA..A.....G...TCA...				
3b Tr	..C...G...T.....-A-.....C...A.....G...TCA...				
3b NE137	..C...G.....-A-T.....C...A.....G...TCA...				
10a JK049	..C...G...TT.....-A-T.....C...A.....G...TCA...				
4a ED43	..C...G...T.....-T.....C...A.....G.....				
4a EG13	..C...G...T.....-T.....C...A.....G.....				
5a SA1	.....G...T.....-A.....C.....G.....				
5a SA7	.....G...T.....-A.....C.....G.....				
FD 2	.....G...T.....-A.....				
6a HK4	.....CA.....CA.....				
6a T3950	.....T.....CA.....CA.....				
11a JK046	.....-.....G.....				

Fig. 1. Comparison of 5' NCR nucleotide sequences obtained from the discrepant samples with sequences specific for each genotype.

type-specific primers (GD2 and GD8). Sequencing results from the remaining haemophilic discrepant sera were based on analysis of single clones only.

## RESULTS

### HCV Positive Haemophiliacs

From a group of 78 HCV PCR positive German haemophilic samples tested with both the Inno-LiPA II assay and the Murex 1-6 serotyping assay, a total of eleven gave discordant results (Table I). In eight cases the genotype identified in the 5' NCR was HCV-3a, with type 1 specific antibody being detected by the ELISA. Two samples (GD10 and GD11) were type 2b by Inno-LiPA and type 1 by serotyping, while GD6 was the opposite, being typed as 1b by Inno-LiPA II and as type 2 by serotyping. In sample GD1, types 1 and 2 were detected by serotyping, but the genotype was HCV-3a by Inno-LiPA.

RNA sequences from the eleven samples were amplified by PCR in the 5' NCR and NS4 regions, and where serum volume permitted, in the core region. The HCV genotype identified in each sample by direct sequence analysis of the 5'NCR [Smith et al., 1995] was concordant with that detected by the Inno-LiPA II assay (Fig. 1), and in each case phylogenetic analysis of sequences in both the core and NS4 regions gave the same genotype assignments as deduced for the 5' NCR. In order to rule out the possibility of a dual infection, PCR amplification of NS4 was carried out using primers specific for the genotypes detected by Inno-LiPA and serotyping (types 1, 2, or 3). In all but two samples, a positive PCR result was identified only when we were using primers specific to the genotype identified by Inno-LiPA (Table I). However, NS4 sequences could be amplified from samples GD2 and GD8 by both type 1 and type 3 specific primers, although the PCR products

TABLE I. Identification of Genotypes in Discrepant Sera Using Different Typing Assays and Sequencing

Sample	Typing results			Type-specific PCR			Sequencing results		
	Inno-LiPA II	Murex 1-3	Murex 1-6	T1	T2	T3	5'NCR	Core	NS4
GD1	3a	—	1 + 2	—		+	3a	—	3a
GD2	3a	—	1	+w		+	3a	3a	3a
GD3	3a	—	1	—		+	3a	3a	3a
GD4	3a	—	1	—		+	3a	3a	3a
GD5	3a	—	1	—		+	3a	—	3a
GD6	1b	—	2	+	—		1b	1b	1b
GD7	3a	—	1	—		+	3a	—	3a
GD8	3a	—	1	+w		+	3a	3a	3a
GD9	3a	—	1	—		—	3a	—	—
GD10	2b	—	1	—	+		2b	2b	2b
GD11	2b	—	1	—	+		2b	2b	—
	Inno-LiPA I								
FD1	1b	NR	4				1b	1b	1b
FD2	4-5	1	1				5	5	5
FD3	3a	NR	1				3a	3a	3a
FD4	3a	NR	1				3a	3a	3a

w = weak result.

with type 1 primers were much weaker under U.V. transillumination than the corresponding bands on the gel using type 3 primers. Multiple clones from each PCR product from GD2 and GD8 were analysed by sequencing, and each was identified as type 3.

One potential explanation for the cases of mistyping by the serotyping assay is that amino acid variation within NS4 in the discrepant samples was sufficient to produce cross-reactivity with peptides normally specific to a different genotype. However, few amino acid differences were observed from the peptides used for serotyping, and were generally restricted to substitutions of residues with similar size and charge (Fig. 2). Although the analysis of multiple clones revealed the same genotype for each isolate, where amino acid sequences from multiple clones show variation, each sequence is shown for comparison.

### French Patients With Chronic Hepatitis C

A total of 88 samples from patients with chronic hepatitis C infection in France were analysed using Inno-LiPA I, and the Murex 1-6 serotyping assay. Four samples were discrepant; sample FD1 was typed as HCV-1b by Inno-LiPA I and yet only had evidence of type 4 specific antibody upon testing by the ELISA, FD2 was type 4 or 5 by Inno-LiPA I, but HCV-1 by serotyping, while samples FD3 and FD4 were HCV-3a by Inno-LiPA I but type 1 by serotyping (Table I).

Sequence analysis in 5'NCR of FD1, FD3, and FD4 confirmed the genotype as that identified by Inno-LiPA I, while sample FD2, identified as 4-5 in the 5'NCR by genotyping, was type 5 upon sequencing in all three regions (Fig. 2).

Amino acid substitutions in NS4 were more frequent than in the haemophiliac samples, and occurred in both peptide regions 1 and 2 (Fig. 2). Two amino acid substitutions occurring in region 2 of sample FD1 pro-

duced a sequence that resembled that of the peptide of HCV-4a used in the serotyping assay.

### DISCUSSION

Genotyping assays are usually based upon the amplification of HCV sequences in the 5'NCR. These allow the identification of each of the six major genotypes but are more limited in their ability to differentiate between subtypes, such as type 1a and 1b, or between type 2a and 2c, which are often identical in this region [Stuyver et al., 1995; Smith et al., 1995]. In addition, the use of genotyping assays presupposes access to PCR technology in the laboratory, and until automation becomes widely available, such methods will be relatively time consuming and expensive. Serotyping assays based on antigenic variation of NS4 or core can distinguish between virus types but are unable to differentiate between subtypes [Machida et al., 1992; Bhattacharjee et al., 1995; Dixit et al., 1995; Zhang et al., 1995]. However, the assays are generally quick and easy to perform and use equipment and techniques that are commonplace in any diagnostic virology laboratory.

A high concordance has been found between the genotypes identified by PCR and the detection of genotype-specific antibody. For example, among reactive samples from Scottish blood donors to branched peptides from NS4, the antibody specificity was 99.2% concordant with the genotypes identified by restriction fragment length polymorphism (RFLP) [Simmonds et al., 1993b]. Similarly, 98.5% of samples were concordant with genotyping by Inno-LiPA I [Lau et al., 1996]. Serologically typeable samples in the modified assay that extended serological detection of types 4, 5, and 6 showed a similar concordance with RFLP (100%, Navas et al., 1997; 98%, Bell et al., 1996; 97%, Bhattacharjee et al., 1995) and other PCR-based typing methods

	1680	1700	1720	1740
Prototype	AAAYCLSTGCV VIVGRVVLGSGKPAIIPDREVLYREFDEMEECSQHLPHYIEQGMMLAEQFKQ KALGLLQTAS RQAEVIAPAV QTNWQKLETF			
Genotype 1 peptides	.....A.....			
	.....V.....			
	R..VV.....Q.....			
	R..V.....Q.....			
1a HCV-PT	.....			
1a HC-J1	.....I.....R.....T.....A.....			
1b HCV-J	.....T..S. ....II...R..V.....Q.....AS.....Q.....T K...AA..V. ESK.RA..V.			
1b BK	.....T..S. ....II...R..V.....L.Q.....AS.....Q.....T K...AA..V. ESK.RA....			
GD 6	.....T..S. ....II.....V.....Q.....AS.....Q.....T K...AA..V. ESK..A..A			
FD 1	.....T..S. ....II.....QA.....AS.....H..Q.....T K...AA..V. ESK.RA....			
1c HC-J9	.....S. ....II.....V.....AA..I..L..H.....K...T.T...H.....S.			
Genotype 2 peptides	RAV.A..K....EA.... ..ASKAAL..E.QRM..ML			
	RVVVV..K..I..EA.... ..ASRAAL..E.QRI..ML			
	RTV.A..K....EA.... ..ASRTAL..E.HRR..ML			
	RAVVA..K....EA....			
2a HC-J6	.....A.... C.I..LHVNRRAVVA..K....EA.....ASRAAL..E.QRI..ML.S ..IQ....Q.. K..QD.Q... ..AS.P.V.Q.			
2a D11353	.....A.... S.I..LHINQRAV.A..K....EA.....ASKATL..E.QRI..ML.S ..IQ....Q.. K..QD.Q... ..S.P.V.Q.			
2b HC-J8	.....A...I S.I..LH.NDRVVVA..K..I..EA.....ASKAAL..E.QRM..ML.S ..IQ....Q.T ..QD.Q..I ..SS.P...Q.			
2b T59	.....A...I S.I..LH.NDRVVVT..K..I..EA.....ASKAAL..E.QRM..ML.S ..IQ....Q.T ..QD.Q.V.. ..SS.P...Q.			
GD 10	.....A...I S.I..LH.NDQVI.A..K..I..EA.....ASKAAL..E.TRM..ML.S ..IQ....Q.T E..QDMQ..I RSS.P...Q.			
GD 11.2	.....A.... S.I..LH.NDQVVA..K..I..EA.....ASKAAL..E.HRI..ML.S ..VQ....Q.T ..QD.Q..I ..SS.P...Q.			
2c BEBE1	.....A.... S.I..LHVNQ..TI.A..K....EA.....ASRTAL..E.HRI..ML.S ..IQ..M.Q.. K..QGVQ... ..AT.P...Q.			
2c T983	.....A.... S.I..GIH.NQRTV.A..K....EA.....ASRTAL..E.HRR..ML.S ..IQ..M.Q.. K..QD.Q.V.. ..GT....Q.			
Genotype 3 peptides	...LV..K....QQY... ..AA.....AQVI..H..			
	.....			
3a NZL1	.....V.... ..HIE.E....LV..K....QQY.....AA.....AQVI..H..E ..I.....R.T Q.QA..E.I..T.....A.			
3a HPCMK6	.....V.... ..HIE.G....LV..K....QQY.....AR.....AQVI..H..E ..V.....R.T Q.QA..E.I..VS.....VL			
GD 1.1	.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..VS...N..VL			
GD 1.3	.....V.... ..IE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.Q			
GD 1.5	SG.....V.... ..HIE.G....LV.GK....QQY.....AA.....AQVI..H..E ..VF.....R.T Q.QQAVIEPI VV			
GD 2A.5	.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVM..H..E ..V.....R.T Q.QA..E.I..AS..K...V.			
GD 2B.5	.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVM..H..E ..V.....R.T Q.QA..E.I..AS..K...V.			
GD 3.4	SS.....E.... ..HIE.G....LV..K....QQY.....AA.....AQAI..H..E ..SF.....R.T Q.QA..E.I..T...R...A.			
GD 3.5	.....E.S. ....HIE.G....LV..K....QQY.....AA.....AQAI..H..E ..I.....R.T Q.QA..E.I..T.....A.			
GD 3.6	.....A.... ..HIE.G....LV..K....QQY.....AA.....AQAI..H..E ..I.....R.T Q.QA..E.I..A.....			
GD 4.2	.....V.... ..HTE.G.R..LV..K....QQY.....AA.....AQAI..H..E ..I.....R.T Q.QA..E.I..A.....A.			
GD 4.3	.....V.... ..HIE.G....LV..K....QQY.....AA.....AQAI..H..E ..I.....R.T Q.QA..E.I..V.....A.			
GD 5.1	SG.....V.... E....HIE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..VS.....VS			
GD 5.2	SG.....V.... ..HIE.G....LV..K....QQY.....AA.....AQV..H..E ..V.....R.T Q.QA..E.I..VS.....A.			
GD 5.3	SG.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..VS.....V.			
GD 7	GH.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..V.....			
GD 8A.2	RP.....V.... ..HIE.G....LV..K....QQY.....G....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..VS.....			
GD 8A.4	S.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..VS.....AL			
GD 8B.5	S.....V.... ..HIE.G....LV..K....QQY.....AAL.....AQVI..H..E ..V.....R.T Q.QA..E.I..TS.....AL			
GD 8B.6	S.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..VS.....AL			
FD 3	.....V.... ..HIE.G....LV..K....QQY.....AA.....AQAI..H..E ..I.....R.T Q.QA..E.I..A.....A.			
FD 4	G.....V.... ..HIE.G....LV..K....QQY.....AA.....AQVI..H..E ..V.....R.T Q.QA..E.I..A.....			
3b Tr	.....V.... ..HIE.G....LV...Q....QQY.....SA.....AQAI..Q...D ..V.....R.. Q.EAE.R.I..SQ...A.A.			
10a JK049	T....V.S. ....HL..G....LV..K....QQY.....RAA.....AQGI..Q...E ..VI....Q.D QK.AD.K.IA TPY.....			
Genotype 4 peptides	Q..V.....QQ.... ..K...LV.H.LQ.....			
	.....			
4a ED43	.....V.S. ....Q..V.....QQ.....K...LV.H.LQ.....NF.G K..QEAT.VI ..S.FA...Q.			
4a EG13	.....V.S. ....Q..V.....QQ.....K...LV.H.LQ.....V...NL.G K..QEAT.VI ..S.FA...Q.			
Genotype 5 peptides	R.....QQ.... ..TS...MDEARAI.G..			
	.....			
5a T478	.....TV.S. T....II...R.....QQ..K....TS...MDEARAI.G...D ..V...IG..G QK...TLK..A TSM.NRA.Q.			
5a SC6	.....TV.S. G....II...R..V.....QQL.....AS...MDEARAI.G...E ..M...IG..G Q...TLK..A TSM.H.V.Q.			
5a SC23	.....STV.S. A....I...R.....QQ.....TS...MDEARAI.R...E ..V...IG..G QK...TLK..A TSM.NRA.Q.			
5a SC24	.....TV.S. A....II...R.....QQ.....TS...MDEARAI.G...E ..VPR..IG..G Q...TLK..A TSM.H.V.Q.			
FD 2	G....TM.S. A....II...V.....QQ.....TS...MGEARAI.G...E ..V...IG..G EK...TLK..A TSM.NRA.Q.			
Genotype 6 peptides	...VV.....I..QQ.... ..R..I..LAE.QQI....			
	.....			
6a HK4	.....V.... ..C...IT.T...VV...I..QQ.....K..I..LAE.QQI....R..V.....ASA K...ELK...HSA.P.V.E.			
6a T3950	.....V.... ..C...T.T...VV...I..QQ.....R..I..LAE.QQI....R..V.....ASA K...ELK...HSA.P.M.E.			
11a JK046	.....TV.S. ..C...ITT.SR..V.....M.QQY.....R...LVE.QQ.....NV...I..VTT K...ELK...HSA.P...Q.			

Fig. 2. Comparison of NS4 amino acid sequences obtained from the discrepant samples with those of peptides used in the serotyping assay. Peptides used in the serotyping assay are in boxes. “.” indicates sequence identity to HCV-PT. Samples identified with a decimal point represent the reference number of the clone (e.g., GD 3.6 represents clone 6 from GD 3). GD 2A, 8A: sample GD2 and GD8 amplified with type 1 specific primers. GD 2B, 8B: sample GD2 and GD8 amplified with type 3 specific primers.



such as Inno-LiPA, of 94% [Vandoorn et al., 1996]. Other serological typing assays involving the detection of type specific antibody to the core protein have also been shown to be concordant with PCR based typing methods (99%, Dixit et al., 1995 and 100%, Machida et al., 1992), although the genotypes that were detected by these assays was restricted to types 1 and 2.

Discrepancies between the results of serotyping by Inno-LiPA and by the serotyping assay could arise because the serotyping assay detects circulating antibody rather than the infecting virus, and so discrepant results could result from antibody induced by previous infections, especially for samples from individuals with a history of multiple exposure, such as haemophiliacs. Alternatively, the Inno-LiPA and serotyping assays could produce discrepant results if an individual were infected with a recombinant virus that contained 5'NCR and NS4 sequences derived from viruses of different genotypes. Finally, the discordant results could be due to mistyping by one or other of the assays. Unusual nucleotide sequence variation in the 5' NCR could produce inappropriate hybridisation of genotype-specific probes. Alternatively, amino acid changes in NS4 could influence the adsorption of antibody to peptides normally specific for a heterologous genotype and produce an incorrect result in the serotyping assay.

### Multiple Infection

A higher frequency of discrepant results between the two assays was observed in haemophiliacs, who have been multiply exposed to a range of HCV variants (14.1%), than in a patient group where multiple exposure would be expected to be less common (4.5%). Sequence analysis of discrepant samples revealed that the sequence of the 5'NCR was always consistent with the genotype of the predominant circulating variant identified by Inno-LiPA. Furthermore, in all discrepant samples, both the core and NS4 sequences were consistent with the genotype identified by sequencing of the 5'NCR. Type-specific primers in NS4 generally amplified the genotype detected by Inno-LiPA, and in the two cases where amplification was obtained using both the type 1 and 3 primers, only type 3 sequences were detected among the clones derived from the PCR products. Overall, we found no evidence for concurrent infection with multiple genotypes in this risk group. However, there was also no evidence for that the serotyping assay incorrectly identified the specificity of the antibody response since the NS4 amino acid sequence of virus from the discrepant samples corresponded closely to that of the corresponding type-specific peptide in both regions 1 and 2 (Fig. 2). Substitutions occurring in these regions were generally to residues of similar size and charge to those in the peptides (an example was the substitution of lysine to arginine in region 1 of GD 4.2). Another conservative amino acid substitution occurred at position 1723 of type 3a viruses, where valine was often replaced with alanine. Only GD 11 showed an obviously nonconservative amino acid substitution that could affect the antigenic-

ity of NS4a, occurring at the N-terminal end of the region 1 peptide.

For the above reasons, the most likely explanation for the observed discrepant results is multiple infection, particularly in the haemophiliacs who have been exposed to a large number of HCV-contaminated blood products in the past. For example, if a patient initially infected with type 1 was recently reinfected with type 3, only antibody specific for type 1 would be detected within the "window period" for the superinfecting virus, and a discrepant result would be obtained. However, since 1985, all factor VIII and IX concentrates have undergone virus inactivation treatment, so recent *de novo* infection in this patient group can be ruled out. On the other hand, we have previously found that the major circulating genotype changed in around 30% of haemophiliacs over the seven to ten years during treatment with exclusively heat-inactivated clotting factor. In the absence of re-exposure to HCV, these changes in genotype must have resulted in reactivation rather than reinfection, which indicates that variants of HCV may persist for several years but can remain undetectable by PCR of plasma [Jarvis et al., 1994]. In some haemophiliacs, changes in circulating genotype were followed by corresponding changes in type-specific antibody, whereas in others, no change in antibody specificity developed, and the delays or absence of the appearance of antibody to the new genotype in this study group led to several discrepant results between the serotyping and genotyping assays. Multiple infection is also a possible explanation for the discrepant results observed in FD1, FD2, and FD3, as their risk factors for infection were intravenous drug abuse, an activity likely also to lead to multiple exposure and infection with more than one genotype. Without a comprehensive analysis of genotypes present in the liver or possible extrahepatic sites of HCV replication, we therefore cannot conclude that genotypes detected in plasma are necessarily the only replicating variants of HCV *in vivo*. Further work characterising the variants in different organs targeted by HCV is clearly necessary to further explore the significance of the discrepant results.

### Recombination

There are some reports describing recombinants of HCV [Yun et al., 1996; Kato et al., 1992], although it is difficult to exclude the possibility that these observations resulted from mixed infection, PCR contamination or artifactual recombination during PCR [Meyers et al., 1990]. Indeed, a consistent feature of other studies has been that sequence relationships between subgenomic regions always reflect those of the complete genome [Ohba et al., 1995; Smith et al., 1995; Simmonds et al., 1994; Stuyver et al., 1994]. In this study, the same genotype was detected by sequencing 5'NCR, core, and NS4 regions, including the NS4 sequences amplified using type-specific primers. Recombination was therefore not the explanation for any of the discrepant results observed in this study.

### Antigenic Variation of NS4

Heterogeneity of the deduced NS4 amino acid sequences in the regions targeted by the serotyping assay was greater in two samples from the French hepatitis C study group than for any from the German haemophiliacs. Sample FD 1 (Inno-LiPA type 1b; serotyping assay type 4) had amino acid substitutions in both peptide regions 1 and 2 (Fig. 1), including the substitution of glutamine for histidine at position 1720. This substitution is also present in the type 4 peptide and is consistent with the observed cross-reactivity in the serotyping assay.

Sequence differences from the type 5a peptide were also observed in sample FD 2 (Inno-LiPA type 4-5, serotyping type 1). In region 1 (positions 1691 and 1694) the sequence of FD2 corresponded closely to one of the type 1 peptides used in the assay (Fig. 2), while in region 2, a substitution of aspartate to glycine (position 1719) was observed, although the sequence remained quite distinct from the region 2 peptides of type 1 (Fig. 2).

In conclusion, the reliability of serological typing methods may be influenced by the patient groups under investigation. Methods based on direct virus detection (such as Inno-LiPA) may be more appropriate for use with patients previously known to have been multiply exposed, such as haemophiliacs, since circulating antibody from previous or multiple infection may occur in these individuals. However, the frequency of discrepant results within samples from patients with chronic disease was much lower, and the simplicity and rapidity of the serotyping assay suggests that it may be a suitable alternative in such study groups.

### REFERENCES

- Bell H, Hellum K, Harthug S, Maeland A, Ritland S, Myrvang B, Vonderlippe B, Raknerud N, Skaug K, Prescott L, Simmonds P, Bucher A, Lind E, Schulz T, Sundoy A, Barstad S, Storset S, Nordoy I, Bjark P, Anfinson OG, Klem K, Stray N, Weberg R, Melsom M, Mosvold J, Haug J, Langtind J, Wetterhus S, Ostborg J, Aadland E, Mowe M, Jahnsen J, Muller F (1996): Prevalence of hepatitis C genotypes among patients with chronic hepatitis C in Norway. *Scandinavian Journal of Infectious Diseases* 28:357-359.
- Bhattacharjee V, Prescott LE, Pike I, Rodgers B, Bell H, Elzayadi AR, Kew MC, Conradie J, Lin CK, Marsden H, Saeed AA, Parker D, Yap PL, Simmonds P (1995): Use of NS-4 peptides to identify type-specific antibody to hepatitis C virus genotypes 1, 2, 3, 4, 5 and 6. *Journal of General Virology* 76:1737-1748.
- Bukh J, Purcell RH, Miller RH (1993): At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. *Proceedings of the National Academy of Sciences (USA)* 90:8234-8238.
- Chan S-W, McOmish F, Holmes EC, Dow B, Peutherer JF, Follett E, Yap PL, Simmonds P (1992): Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants. *Journal of General Virology* 73:1131-1141.
- Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M (1989): Isolation of a cDNA derived from a blood-borne non-A, non-B hepatitis genome. *Science* 244:359-362.
- Davidson F, Simmonds P, Ferguson JC, Jarvis LM, Dow BC, Follett EAC, Seed CRG, Krusius T, Lin C, Medgyesi GA, Kiyokawa H, Olim G, Duraisamy G, Cuypers T, Saeed AA, Teo D, Conradie J, Kew MC, Lin M, Nuchaprayoon C, Ndimbie OK, Yap PL (1995): Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' noncoding region. *Journal of General Virology* 76:1197-1204.
- Davis GL (1994): Prediction of response to interferon treatment of chronic hepatitis C. *Journal of Hepatology* 21:1-3.
- Dixit V, Quan S, Martin P, Larson D, Brezina M, Dinello R, Sra K, Lau JYN, Chien D, Kolberg J, Tagger A, Davis G, Polito A, Gitnick G (1995): Evaluation of a novel serotyping system for hepatitis C virus: Strong correlation with standard genotyping methodologies. *Journal of Clinical Microbiology* 33:2978-2983.
- Dusheiko GM, Simmonds P (1994): Sequence variability of hepatitis C virus and its clinical relevance. *Journal of Viral Hepatitis* 1:1-13.
- Jarvis LM, Watson HG, McOmish F, Peutherer JF, Ludlam CA, Simmonds P (1994): Frequent reinfection and reactivation of hepatitis C virus genotypes in multitransfused hemophiliacs. *Journal of Infectious Diseases* 170:1018-1022.
- Kato N, Ootsuyama Y, Tanaka T, Nakagawa M, Nakazawa T, Muraiso K, Ohkoshi S, Hijikata M, Shimotohno K (1992): Marked sequence diversity in the putative envelope proteins of hepatitis C viruses. *Virus Research* 22:107-123.
- Kleter GEM, Van Doorn LJ, Brouwer JT, Schalm SW, Heijntik RA, Quint WGV (1994): Sequence analysis of the 5' untranslated region in isolates of at least four genotypes of hepatitis C virus in the Netherlands. *Journal of Clinical Microbiology* 33:306-310.
- Kobayashi M, Tanaka E, Sodeyama T, Urushihara A, Matsumoto A, Kiyosawa K (1996): The natural course of chronic hepatitis C: A comparison between patients with genotypes 1 and 2 hepatitis C viruses. *Hepatology* 23:695-699.
- Lau JYN, Davis GL, Prescott LE, Maertens G, Lindsay KL, Qian KP, Mizokami M, Simmonds P, Perrillo RP, Schiff ER, Bodenheimer HC, Balart LA, Regenstein F, Dienstag JL, Katkov WN, Tamburro CH, Goff JS, Everson GT, Goodman Z, Albrecht J (1996): Distribution of hepatitis C virus genotypes determined by line probe assay in patients with chronic hepatitis C seen at tertiary referral centers in the United States. *Annals of Internal Medicine* 124:868.
- Machida A, Ohnuma H, Tsuda F, Muneakata E, Tanaka T, Akahane Y, Okamoto H, Mishiro S (1992): Two distinct subtypes of hepatitis C virus defined by antibodies directed to the putative core protein. *Hepatology* 16:886-891.
- McOmish F, Yap PL, Dow BC, Follett EAC, Seed C, Keller AJ, Cobain TJ, Krusius T, Kolho E, Naukkarinen R, Lin C, Lai C, Leong S, Medgyesi GA, Hejjas M, Kiyokawa H, Fukada K, Cuypers T, Saeed AA, Alrasheed AM, Lin M, Simmonds P (1994): Geographical distribution of hepatitis C virus genotypes in blood donors—an international collaborative survey. *Journal of Clinical Microbiology* 32:884-892.
- Mellor J, Holmes EC, Jarvis LM, Yap PL, Simmonds P, International Collaborators (1995): Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification. *Journal of General Virology* 76:2493-2507.
- Mellor J, Walsh EA, Prescott LE, Jarvis LM, Davidson F, Yap PL, Simmonds P, Nowicki MJ, Mosley JW, Lin CK, Lai CL, deOlim G, Martins IA, Ong YW, Teo D, Lin M, Nuchaprayoon C, Tanprasert S (1996): Survey of type 6 group variants of hepatitis C virus in southeast Asia by using a core-based genotyping assay. *Journal of Clinical Microbiology* 34:417-423.
- Meyerhans A, Vartanian JP, Wain Hobson S (1990): DNA recombination during PCR. *Nucleic Acids Research* 18:1687-1691.
- Navas S, Castillo I, Martin J, Quiroga JA, Bartolome J, Carreno V (1997): Concordance of hepatitis C virus typing methods based on restriction fragment length polymorphism analysis in 5' noncoding region and NS4 serotyping, but not in core PCR or a line probe assay. *J. Clin. Micro.* 35:317-321.
- Ohba K, Mizokami M, Ohno T, Suzuki K, Orito E, Ina Y, Lau JYN, Gojbori T (1995): Classification of hepatitis C virus into major types and subtypes based on molecular evolutionary analysis. *Virus Research* 36:201-214.
- Ohno T, Mizokami M, Wu RR, Saleh MG, Ohba KI, Orito E, Mukaide M, Williams R, Lau JYN (1997): New Hepatitis C Virus (HCV) Genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. *Journal of Clinical Microbiology* 35:201-207.
- Okamoto H, Sugiyama Y, Okada S, Kurai K, Akahane Y, Sugai Y, Tanaka T, Sato K, Tsuda F, Miyakawa Y, Mayumi M (1992): Typing hepatitis C virus by polymerase chain reaction with type-specific primers: application to clinical surveys and tracing infectious sources. *Journal of General Virology* 73:673-679.
- Okamoto H, Tokita H, Sakamoto M, Horikita M, Kojima M, Iizuka H, Mishiro S (1993): Characterization of the genomic sequence of type

- V (or 3a) hepatitis C virus isolates and PCR primers for specific detection. *Journal of General Virology* 74:2385–2390.
- Prescott LE, Simmonds P, Lai CL, Chan NK, Pike I, Yop PL (1996): Detection and clinical features of hepatitis C virus type 6 infections in blood donors from Hong Kong. *Journal of Medical Virology* 50(2):168–175.
- Simmonds P, Holmes EC, Cha TA, Chan S-W, McOmish F, Irvine B, Beall E, Yap PL, Kolberg J, Urdea MS (1993a): Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *Journal of General Virology* 74:2391–2399.
- Simmonds P, Rose KA, Graham S, Chan SW, McOmish F, Dow BC, Follett EAC, Yap PL, Marsden H (1993b): Mapping of serotype-specific, immunodominant epitopes in the NS-4 region of hepatitis C virus (HCV)—use of type-specific peptides to serologically differentiate infections with HCV type 1, type 2, and type 3. *Journal of Clinical Microbiology* 31:1493–1503.
- Simmonds P, Smith DB, McOmish F, Yap PL, Kolberg J, Urdea MS, Holmes EC (1994): Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1, and NS-5 regions. *Journal of General Virology* 75:1053–1061.
- Smith DB, Mellor J, Jarvis LM, Davidson F, Kolberg J, Urdea M, Yap PL, Simmonds P, Conradie JD, Neill AGS, Dusheiko GM, Kew MC, Crookes R, Koshy A, Lin CK, Lai C, Murraylyon IM, Elguneid A, Gunaid AA, Yemen T, Yemen S, Mutimer D, Ahmed M, Nuchprayoon C, Tanprasert S, Preston FE, Makris M, Chuansumrit A, Mahasandana C, Pritchard D, Riley E, Greenwood BM, Saeed AA, Alrasheed AM, Saleh MG, Mcfarlane I, Tibbs C, Williams R, Power J, Lawlor E, Kiyokawa H (1995): Variation of the hepatitis C virus 5' noncoding region: Implications for secondary structure, virus detection and typing. *Journal of General Virology* 76:1749–1761.
- Stuyver L, Rossau R, Wyseur A, Duhamel M, Vanderborgh B, Van Heuverswyn H, Maertens G (1993): Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. *Journal of General Virology* 74:1093–1102.
- Stuyver L, Vanarnhem W, Wyseur A, Hernandez F, Delaporte E, Maertens G (1994): Classification of hepatitis C viruses based on phylogenetic analysis of the envelope 1 and nonstructural 5b regions and identification of five additional subtypes. *Proceedings of the National Academy of Sciences (USA)* 91:10134–10138.
- Stuyver L, Wyseur A, Vanarnhem W, Lunel F, Laurentpuig P, Pawlotsky JM, Kleter B, Bassit L, Nkengasong J, Vandoorn LJ, Maertens G (1995): Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable. *Virus Research* 38:137–157.
- Stuyver L, Wyseur A, Vanarnhem W, Hernandez F, Maertens G (1996): Second-generation line probe assay for hepatitis C virus genotyping. *Journal of Clinical Microbiology* 34:2259–2266.
- Takada A, Tsutsumi M, Zhang SC, Okanoue T, Matsushima T, Fujiyama S, Komatsu M (1996): Relationship between hepatocellular carcinoma and subtypes of hepatitis C virus: A nationwide analysis. *Journal of Gastroenterology Hepatology* 11:166–169.
- Tanaka T, Tsukiyamakohara K, Yamaguchi K, Yagi S, Tanaka S, Hasegawa A, Ohta Y, Hattori N, Kohara M (1994): Significance of specific antibody assay for genotyping of hepatitis C virus. *Hepatology* 19:1347–1353.
- Vandoorn LJ, Kleter B, Pike I, Quint W (1996): Analysis of hepatitis C virus isolates by serotyping and genotyping. *Journal of Clinical Microbiology* 34:1784–1787.
- Widell A, Shev S, Mansson S, Zhang YY, Foberg U, Norkrans G, Fryden A, Weiland O, Kurkus J, Nordenfelt E (1994): Genotyping of hepatitis C virus isolates by a modified polymerase chain reaction assay using type specific primers: epidemiological applications. *Journal of Medical Virology* 44:272–279.
- Yun AB, Lara C, Johansson B, deRivera IL, Sonnerborg A (1996): Discrepancy of hepatitis C virus genotypes as determined by phylogenetic analysis of partial NS5 and core sequences. *Journal of Medical Virology* 49:155–160.
- Zhang ZX, Yun ZB, Chen M, Sonnerborg A, Sallberg M (1995): Evaluation of a multiple peptide assay for typing of antibodies to the hepatitis C virus: Relation to genomic typing by the polymerase chain reaction. *Journal of Medical Virology* 45:50–55.